4Л

УДК 621.307

О.Ю. Ломакіна
Національний технічний університет України «Київський політехнічний інститут»,
м. Київ, Україна
leleka@rambler.ru

Аналіз роботи систем розпізнавання жестів

У наведеній статті розглядається один з перспективних напрямків розвитку інформаційних технологій в області інтерфейсу людина-машина на основі використання природного для людини способу спілкування з комп'ютерами. Надано огляд основних етапів, процедур та технологій збору даних, що застосовуються у роботі системи розпізнавання жестів. Проводиться поетапний аналіз роботи систем розпізнавання жестів, визначені основні проблеми їх побудови та пропонуються оптимальні шляхи їх подальшого вирішення.

Вступ

Перспективним напрямком розвитку інформаційних технологій є розробка нових способів забезпечення інтерфейсу людина-машина. Перед розробниками подібних інтерфейсів ставиться задача використання природних для людини способів спілкування з комп'ютерами. Враховуючи усі можливі перешкоди та наявність шумів в оточуючому середовищі, перевага надається системам на основі комп'ютерного зору. Особливо перспективними для побудови інтерфейсів управління програмним та апаратним забезпеченням комп'ютерів є жести [1-3]. Перевага такого інтерфейсу полягає у тому, що жести дозволяють розширити можливості інтерфейсу для людей з вадами слуху і мови, та забезпечити дистанційне управління різними побутовими пристроями, хоча при їх створенні існує ряд не повністю вирішених задач, які розглядаються у статті.

Розглянемо роботу систем управління жестами, які можуть бути вбудовані у мультимедійні пристрої і заміняти собою пульт дистанційного керування.

Робота системи складається з основних етапів:

1) отримання зображення, збір даних – об'єкт (рука) фіксується за допомогою камер, які підключені до ПК;

2) локалізація руки на зображенні, виділення пальців руки – на отриманому зображенні знаходиться ділянка руки;

3) розпізнавання жестів – виділення характерних ознак (контур руки, колір шкіри);

4) класифікація жестів та присвоєння розпізнаному жесту певної команди.

Метою даної статті є аналіз роботи систем розпізнавання жестів, які розглядаються як системи машинного зору для виявлення та позиціонування жестів людини.

Огляд основних етапів та процедур в системі розпізнавання жестів

Першим етапом будь-якої системи розпізнавання є *збір даних* (як дані може використовуватись інформація про колір).

Інформація, отримана за допомогою зображення, обробляється для виділення області руки від фону. Цей етап називається *локалізацією руки* на зображенні.

Після збору даних стає можливим використання первинної інформації про руку для *фільтрації даних* та видалення шумів на зображенні, які можуть виникнути, наприклад, через перепади освітлення, а також видалення артефактів (наявність на руці каблучок). Дана процедура є надзвичайно важливою, так як набором жестів, які необхідно розпізнати, визначається кількість та розмір отриманих даних і для реалізації якісного розпізнавання необхідною умовою є точність (достовірність) цих даних.

На етапі *розпізнавання рухів рук* здійснюють *виділення ознак*. Вибір ознаки є важливою частиною процесу розпізнавання, тому що рухи рук відрізняються розмаїтістю форм, рухів і текстур. Для розпізнавання статичного положення руки застосовують геометричні ознаки, такі як кінчики пальців, напрямок пальців. Але такі ознаки не завжди доступні і не завжди надійні через самозатінення та умови освітлення. В даній статті як ознаки використовуються властивості контуру, які описують жести.

Наступним етапом буде визначення конкретних жестів на основі аналізу відфільтрованих даних, які несуть інформацію про рух руки. З цією метою проводиться процедура *класифікації*. Але перед цим систему необхідно «навчити» реагувати на жести, та адаптувати їх щодо конкретних рухів рук користувача. В результаті навчання система поступово здобуває здатність відповідати потрібними реакціями на певні сукупності зовнішніх впливів, а в результаті адаптації відбувається підстроювання параметрів і структури системи з метою досягнення необхідної якості керування в умовах безперервних змін зовнішніх умов.

Як об'єкти навчання можуть виступати візуальні зображення рук.

Пропонується використовувати набір жестів з алфавіту глухонімих, що розширить можливості інтерфейсу для людей з дефектами слуху.

Технології збору даних

Для збору інформації в основному застосовуються наступні технології:

1. Контактні. Використовується інформаційна рукавичка (data glove – пристрій введення даних у системі віртуальної реальності, що фіксує рухи руки користувача та передає їх до комп'ютера) для захоплення рухів руки. На основі даних, отриманих за допомогою цих пристроїв програмними засобами, формується модель руки.

2. Безконтактні (оптичні методи) – методи комп'ютерного зору. Переваги контактних технологій полягають у високій завадостійкості та можливості здійснювати процеси в режимі реального часу. До недоліків можна віднести необхідність одягання рукавички, що накладає певні обмеження на користувача. Саме тому перевага була надана безконтактним технологіям, які є більш практичними та ефективнішими, і звільняють рух рук користувача від обмежень.

Для того щоб розпізнати жести руки, необхідно зібрати інформацію про руку за допомогою однієї або декількох камер. Відповідно цього існують різні системи, які можна класифікувати наступним чином:

Стереографічна система надає детальну попіксельну інформацію для будь-якої точки в полі зору камер і, таким чином, забезпечує великий об'єм інформації про руку. За допомогою такої системи можна легко відокремити пальці руки на фоні шкіри, оскільки вони знаходяться ближче до камери. Проте для обчислення 3 D даних знадобиться багато операційного часу, що виключає використання алгоритмів у режимі реального часу.

Багатокамерна 2 D система спостереження буде видавати меншу кількість інформації у порівнянні зі стереографічною, але й потребуватиме менше часу на оброб-

ку (за умови, що кількість камер невелика). За допомогою цієї системи два або й більше 2 D зображень руки, які отримали з різних камер, можна об'єднати для розпізнавання жестів. Достатня кількість камер забезпечить повний об'єм інформації для визначення будь-якого жесту.

Однокамерна система буде забезпечувати значно меншу кількість інформації про руку. Деякі особливості, наприклад, палець на фоні шкіри буде дуже важко визначити, оскільки буде відсутня інформація про його товщину. В основному точно може бути визначена тільки інформація про контур. Дані про контур будуть відносно позбавлені від шумів (за умови, що рука успішно відділяється від фону) та на їх обчислення знадобиться невелика кількість часу.

Можливо визначити велику кількість жестів, використовуючи тільки інформацію про контур, окрім цього однокамерна система вносить менше завад, дешевша та має найкращу швидкодію. Хоча система отримує найменше інформації у порівнянні із попередніми системами, цей недолік повністю перекривається перевагами, які були вказані вище. Тому було вирішено використовувати саме однокамерну систему для збору даних.

Дані на виході вибраної однокамерної системи представляють 2 D масив RGB-пікселів, отриманих в різні проміжки часу. Для того щоб визначити інформацію про контур, необхідно відділити пікселі шкіри від пікселів фону на зображенні. Щоб використовувати інші ознаки для забезпечення додаткової інформації про руку, їх також потрібно відокремити від фону (та пікселів шкіри). Для реалізації цього важливо правильно підібрати обладнання та забезпечити певні умови.

Задача виділення пікселів шкіри буде набагато простішою за умов правильного освітлення. Якщо освітлення в полі зору камери постійне, то ефект самозатінення можна звести до мінімуму. Для забезпечення достатньої кількості світла для ПЗЗ камери яскравість теж має бути наперед задана.

Однак, оскільки система розпізнавання жестів розробляється для звичайного споживача, то вимоги до використання спеціального освітлення будуть накладати певні незручності. Отже, надалі будемо вважати, що інформація про руку визначається за умов освітлення стандартної кімнати, хоча це обмежить використання системи за умов іншого навколишнього середовища.

Локалізація області руки

Вибір колірного простору

Колірний простір являє собою модель подання кольору, засновану на використанні колірних координат. Колірний простір будується таким чином, щоб будь-який колір був представлений точкою, яка має певні координати, причому так, щоб одному набору координат відповідав один колір.

Розрізняють RGB та HSL колірні моделі.

Результати детектування за допомогою HSL-моделі показали, що тон є ненадійним параметром для визначення ділянок шкіри на зображенні, оскільки значення тону шкіри та тону фону частково співпадають. Насиченість дає ненабагато кращі результати, а найбільш оптимальним параметром виявилась яскравість. В даному випадку найкращі результати визначення ділянок шкіри на зображенні дає поєднання значень насиченості та яскравості.

Як показали результати, у випадку використання RGB-моделі, області для кожної колірної компоненти легко відокремити. Враховуючи це і той факт, що обчислення RGB-простору не вимагає додаткового часу на перетворення RGB в HSL, превага була надана саме RGB колірному простору.

Згідно з аналізом результатів тестування за різних умов освітлення, яке виконувалось над послідовністю зображень з камери, було визначено, що сегментація, яка проводиться в R (r канал у нормалізованому RGB просторі) і S (S канал HSV колірного простору) каналах, є найбільш стійкою до змін освітлення, а також має найменший відсоток помилок 2-го роду. Це свідчить про доцільність застосування саме цього методу сегментації в R і S каналах різних колірних просторів [4].

Обробка зображень. Відокремлення області зап’ястя та передпліччя

На етапі обробки зображення для подальшого коректного розпізнавання жесту необхідно відокремити пікселі шкіри, які описують кисть та будуть у подальшому використовуватись для визначення жесту, від пікселів шкіри, які відносяться до області зап’ястя та області передпліччя. Оскільки для всіх них значення кольору майже однакові, то для визначення та видалення небажаних пікселів буде використовуватись інформація про розташування пікселів руки та зап’ястя.

За допомогою усереднення значень детектованих пікселів можливо обчислити центроїд (центр ваги) руки та зап’ястя.

Набір положень всіх пікселів шкіри обчислюється за формулою:

$$L=\left\{\vec{x}\mid S\left(r(\vec{x}),g(\vec{x}),b(\vec{x})\right)=\mathbf{1}\right\}.$$

Центроїд руки:

$$\vec{c}_p=\frac{\mathbf{1}}{|L|}\sum_{\vec{x}\in L}\vec{x}.$$

Центроїд зап’ястя:

$$\vec{c}_з=\frac{\mathbf{1}}{|L_з|}\sum_{\vec{x}\in L_з}\vec{x},$$

де $|L|$ – кількість елементів L.

Краї зап’ястя можна знайти шляхом сканування ліній паралельно лінії, яка з’єднує два центроїда. Вектор, з’єднуючий два центроїда, можна визначити як:

$$\vec{c}_{dif}=\left(x_{dif},y_{dif}\right)=\vec{c}_p-\vec{c}_з.$$

Кут повороту руки:

$$\theta_p=\tan^{-1}\left(\frac{y_{dif}}{x_{dif}}\right).$$

Тоді краї зап’ястя обчислюються наступним чином для кожної точки $\vec{p}_1(s_1)$ вздовж лінії:

$$\vec{p}_1(s_1)=\vec{c}_з+s_1\begin{pmatrix}\mathrm{Cos}\left(\theta_з+\frac{\pi}{\mathbf{2}}\right)\\ \mathrm{Sin}\left(\theta_з+\frac{\pi}{\mathbf{2}}\right)\end{pmatrix},$$

де ($-50 \le S_1 \le 50$)

Для кожної функції s_1 обчислюється кількість пікселів зап'ястя $n(s_1)$ вздовж лінії

$$\vec{p}_2(s_1, s_2) = \vec{p}_1 + s_2 \begin{pmatrix} Cos(\theta_p) \\ Sin(\theta_p) \end{pmatrix},$$

де $(-50 \le s_2 \le 50)$

Дві точки, які визначають краї зап'ястя $\vec{b}_{left}$ або (x_{left}, y_{left}) та $\vec{b}_{right}$ або (x_{right}, y_{right}), дорівнюють $\vec{p}_1(s_1)$, коли значення $n(s_1)$ нижче певного порогу. За допомогою порогу можливо виділити значення пікселів, які знаходяться в області зап'ястя.

Радіус області зап'ястя визначається як:

$$r_з = \max\left(\left|\vec{b}_{left} - \vec{c}_з\right|, \left|\vec{b}_{right} - \vec{c}_з\right|\right).$$

Виділення пальців руки

Виявлення та локалізація мінімумів, максимумів руки (екстремумів), тобто кінчиків пальців і проміжків між пальцями, надасть додаткову інформацію при виконанні процедури розпізнавання жестів. Оскільки і мінімуми, і максимуми характеризуються великою кривизною, проводяться експерименти з гістограмою кривизни контуру, тобто розглядається ділянка кривизни контуру в різних масштабах уздовж всієї довжини шляху. 9 послідовних максимумів гістограми кривизни беруться як шукані значення кінчиків пальців. Слід зазначити, що цей метод досить чутливий до розривів контуру і виникають проблеми при визначенні порожнин і зламів в області зап'ястя. Тому було зосереджено увагу на більш надійному альтернативному методові, який полягає у побудові графіка залежності відстані за радіусом від опорної точки (точки відліку) в області зап'ястя. Як опорна точка береться перша точка перетинання головної осі з лінією зап'ястя. В результаті одержуємо послідовність відстаней за радіусом, які представляють мінімуми та максимуми, тобто шукані екстремуми. Значення екстремумів є дуже надійними, оскільки на визначення 5 максимумів (пальців) і 4 мінімумів не впливають шуми на контурі.

Для всіх наступних етапів обробки рук використовується інформація, отримана з набору базисних осей чотирьох пальців (вказівний, середній, безіменний, мізинець). Встановлено, що базисна лінія – це лінія, що проводиться через чотири точки відповідно до методу найменших квадратів або просто з'єднує базисні точки вказівного пальця і мізинця. Ця лінія називається основною лінією руки (pivot line). Основна лінія руки використовується для: по-перше, реєстрації всіх зображень руки відповідно до обраного кута повороту лінії. По-друге, кути повороту пальців обчислюються стосовно основної лінії. Нарешті, орієнтація та розмір основної лінії допомагає визначити великий палець і область зап'ястя.

Після того як визначені екстремуми, за допомогою зазначеного вище методу проводяться відрізки уздовж бічної сторони пальця, починаючи від кінчика до двох сусідніх точок проміжку між пальцями. Вибирається найкоротший відрізок і потім від однієї точки до іншої проводиться крива. Отримана крива у формі серпа чітко розмежовує палець, таким чином можна обчислити його довжину. Однак, для великого пальця операції локалізації трохи відрізняються.

Рух великого пальця ускладнюється тим, що він включає обертання відносно двох різних суглобів (п'ясної фаланги та кістки-трапеції). Складність полягає у тому,

що натягнута шкіра між великим і вказівним пальцем вносить погрішності при визначенні западини між пальцями та при визначенні великого пальця. Відповідно до анатомії вважається, що великий палець такої ж довжини, як і мізинець. Проводиться лінія уздовж головних осей пальця і точка на цій лінії, що перебуває на відстані 120% від розміру мізинця, формує базисну лінію великого пальця. Палець переміщується таким чином, що його базисна лінія збігається з вершиною лінії руки, коли остання відхиляється на 90 градусів за годинниковою стрілкою. Великий палець обертається за остаточним напрямком і повертається на місце.

При виявленні та локалізації пальців руки необхідно враховувати такий момент, як обертання пальців. Пальці обертаються навколо точки між основною фалангою та п'ястковою кісткою (мед.). П'ясткова кістка – це кістяк руки між зап'ястям і п'ятьма пальцями. Цей кістяк складається з 5 довгих трубчастих кісток, які перебувають між кістками зап'ястя та пальців (фаланг). Ці з'єднання трохи нижче лінії, що з'єднує проміжки між пальцями. Тому головна вісь кожного пальця збільшується в довжину на 20% і перевищує відповідну довжину пальців. Сукупність кінцевих точок осей чотирьох пальців (вказівний, середній, безіменний, мізинець) визначає лінію, що залежить від розміру та орієнтації руки.

Головні осі кожного пальця обчислюються за допомогою власної матриці інерційних характеристик. Поточний кут напрямку пальця визначається як $\theta = \boldsymbol{arctan(}v_{maj} \boldsymbol{/} u_{maj}\boldsymbol{)}$, де $\boldsymbol{(}u_{maj}\boldsymbol{,}v_{maj}\boldsymbol{)}$ – основний власний вектор. Кожен палець i повертається на кут $\Delta\theta_i = \theta_i - \psi_i$, при i = вказівний палець, середній, безіменний, мізинець, а ψ_i – напрямок пальця. Обертання пальця здійснюється шляхом множення вектора положення пікселів пальця на матрицю обертання навколо базисної лінії (pivot):

$$R = \begin{bmatrix} \boldsymbol{cos(}\Delta\theta\boldsymbol{)} & \boldsymbol{-sin(}\Delta\theta\boldsymbol{)} \\ \boldsymbol{sin(}\Delta\theta\boldsymbol{)} & \boldsymbol{cos(}\Delta\theta\boldsymbol{)} \end{bmatrix}.$$

Стандартні кути пальців для середньої руки наведені в табл. 1. Будь-який інший набір кутів буде однаково добре працювати в даному алгоритмі за умови, що при цьому наборі кутів пальці будуть відокремлені один від одного.

Таблиця 1 – Кути обертання пальців

Великий палець	Вказівний	Середній	Безіменний	Мізинець
150	120	100	80	60

Розпізнавання жестів

Аналіз існуючих методів

Методи розпізнавання в системах комп'ютерного зору можна розділити на два види – методи на основі створення тривимірної моделі руки й методи на основі виділення ознак [5].

Методи на основі створення тривимірної моделі руки базуються на побудові кінематичної моделі, яка враховує всі можливі ступені свободи [6]. Для цього потрібно оцінити жести руки за допомогою порівняння положення руки на вхідному зображенні та двовимірної проекції моделі жесту з бази даних. Такі методи потенційно

дозволяють розпізнавати значну кількість жестів. Однак, з метою його реалізації потрібне створення великої бази даних зображень для порівняння з побудованою моделлю та подолання складностей при виділенні ознак з урахуванням анатомічних особливостей.

Методи на основі виділення ознак засновані на врахуванні особливостей зображень, які використовуються для визначення положення руки. Одним з підходів є знаходження ділянок шкіри на зображенні з використанням колірних ознак.

Методи на основі виділення ознак застосовуються за умови, якщо можливо ідентифікувати характерні точки або області на об’єктах, а сам об’єкт може бути представлений як сукупність цих областей. У цьому випадку замість моделювання об’єкта в цілому, моделюється деяка множина характерних областей. Перевага цього підходу полягає в тому, що він стійкий до деформацій та інших змін вхідних даних. При визначенні достатньої кількості характерних точок, об’єкт надійно ідентифікується. Для завдань визначення та розпізнавання жестів у реальному часі дані методи є перспективними, тому було вирішено застосовувати саме їх.

Вибір методу розпізнавання

На етапі розпізнавання рухів рук здійснюють виділення ознак. Вибір ознаки є важливою частиною процесу розпізнавання, тому що рухи рук відрізняються різноманітністю форм, рухів і текстур. Для розпізнавання статичного положення руки застосовують геометричні ознаки, такі як кінчики пальців, напрямок пальців. Але такі ознаки не завжди доступні і не завжди надійні через самозатінення та умови освітлення. В даній роботі як характерна ознака використовуються властивості контуру.

Для отримання характерних ознак, а саме інформації про контур, застосовується аналіз незалежних компонент.

Аналіз незалежних компонент

Аналіз незалежних компонент (АНК) використовується для вибору статистично незалежних змінних з їх сукупності [7]. Він успішно застосовується в різних додатках для пошуку схованих коефіцієнтів серед даних для аналізу або ж для їхнього розкладання на вихідні сигнали; для реального зображення АНК також служить корисним інструментом для отримання характерних ознак та завдань ідентифікації людини.

Як вхідні дані використовуються 1-D бінарні сигнали джерела, які несуть інформацію про контур руки на зображенні.

АНК припускає, що кожний зі спостережуваних сигналів $\{x_i(k), k = 1, \ldots, K\}$ являє собою суміш із набору N-невідомих сигналів незалежних джерел s_i в невідомій матриці змішування A. x_i та s_i формують рядки в матрицях X та S, розмірністю N×K, відповідно: X = AS.

Вектори даних для АНК – це пікселі руки на зображенні. K – розмірності цих векторів (наприклад, K = 40000, якщо є зображення руки розміром 200×200). Ціль АНК – знайти таке лінійне перетворення W для вхідних даних, яке мінімізує статистичну залежність між вихідними компонентами y_i, за допомогою яких оцінюються передбачувані незалежні джерела s_i:

$$\hat{S} = Y = WX.$$

Для того щоб знайти W-перетворення, яке називають матрицею розкладання, застосовується швидкий алгоритм АНК: він максимізує статистичну незалежність між вихідними компонентами з використанням максимізації їхньої негативної ентропії. За допомогою АНК можна одержати незалежні базисні зображення або ж незалежні коефіцієнти змішування.

У першому випадку вважається, що N – це дані про руку, що являють собою лінійну суміш невідомого набору N-статистично незалежних джерел сигналів, які несуть інформацію про руку («hand sources»).

Для одержання векторів даних застосовується растрове сканування нормованого зображення рук. Для зображень розміром 200×200, матриця даних X буде N × 40000 мірною, отже, m = 40000. Ця матриця розкладається на N незалежних компонентів джерела $\{\hat{s}_i\}$, які формують ряди вихідної матриці $\hat{S} = WX$. Кожен рядок матриці змішування A (N×N) буде містити вагові коефіцієнти, характерні для даної руки. Ці коефіцієнти показують відносний внесок джерела сигналів, які несуть інформацію про руку, у синтез даного зображення руки. Із цього випливає, що тестове зображення руки x_i i-го рядка матриці A буде являти собою N-мірний вектор ознак.

На етапі розпізнавання, вважаючи, що тестовий набір наслідує ту саму модель синтезу з тими ж незалежними компонентами, нормований тестовий набір зображень руки (1×40000) проектується на набір визначених базисних функцій, а результуючий вектор коефіцієнтів проекції обчислюється за формулою:

$$a_{test} = x_{test}\hat{S}^T\left(\hat{S}\hat{S}^T\right)^{-1}.$$

Нарешті, окрема ознака, що тестується, розпізнається як характерна для руки людини i*, коли значення a_{test} перебуває найближче всього до значення вектора ознаки a_{i*}, а відстань вимірюється за допомогою L1 метрики:

$$i^* = \arg\min_{(i)}\left\{\sum_{j=1}^{N}\left|a_{i,j} - a_{test,j}\right|\right\}.$$

У другому випадку вважаються незалежними коефіцієнти накладення, а не базисні зображення. Такий підхід передбачає, що кожен K-піксель зображення руки одержують із незалежних сумішей випадкових величин, тобто джерел сигналів, які несуть інформацію про піксель («pixel sources»). Виходячи із цього, виконується транспонування матриці даних: X^T. Однак величезна розмірність піксельних векторів (як правило, K>> N) вимагає виконання процедури аналізу головних компонентів (АГК) перед реалізацією АНК.

Власні вектори коваріаційної матриці $C = \frac{1}{N}X^T X$, розмірністю K×K, де кожен ряд матриці X^T центрований, можна обчислити за допомогою власних векторів матриці X X^T набагато меншої розмірності N×N.

Нехай $\{v_1,\ldots,v_M\}$ – власні вектори M-ряду із власними значеннями $\{\lambda_1 \geq \lambda_2 \ldots \geq \lambda_M\}$ матриці X X^T, розмірності N×N. Тоді відповідно до теореми розкладання за сингулярними числами матриці (singular value decomposition, SVD), ортонормовані власні вектори $\{w_1,\ldots,w_M\}$ матриці C, що відповідають $M \leq N$ найбільшим власним значенням $\{\lambda_1,\lambda_2\ldots,\lambda_M\}$, визначаються як $w_j = \frac{1}{\sqrt{\lambda_j}}Xv_j$, $j = 1,..,M$. Після проектування вхідного вектора x на власні w_j, ми одержимо j-е ознаку $y_j = \frac{1}{\sqrt{\lambda_j}}v_j^T Xx^T = R\,x^T$, де R являє собою оператор проектування.

Кількість даних зображення руки зменшується після проекції на невелику кількість M-головних компонентів, і таким чином формується квадратна матриця даних RX^T . Потім RX^T розкладається на джерело та коефіцієнти перемішування

залежно від моделі, стовпці оціненої матриці змішування A (N×N) представляють собою базисні функції (зображення руки). І навпаки, коефіцієнти оціненої матриці джерела сигналів є статистично незалежними.

На етапі розпізнавання тестовий набір зображень рук зменшує свою розмірність за допомогою $R x_{test}^{T}$ і множиться на матрицю розкладання $W = A^{-1}$. Результуючий вектор коефіцієнтів тестового набору зображень рук x_{test} (K×1), визначається за формулою $\hat{p}_{test} = W R x_{test}$, і потім порівнюється із заданими векторами ознак на етапі навчання. Нарешті, зображення руки, що тестується, розпізнається як характерне для руки людини i* за допомогою найближчого вектора ознаки $\hat{p}_{i^*}$, де відстань визначається косинусом кута між ними:

$$i^{*} = \arg \max_{(i)} \left\{ \frac{\hat{p}_i \cdot \hat{p}_{test}}{\|\hat{p}_i\| \|\hat{p}_{test}\|} \right\}.$$

Після розпізнавання жестів відбувається процедура класифікації та присвоєння команди. Види команд визначаються специфікою застосування розробленої системи розпізнавання. Наведені системи управління можуть бути інтегровані в побутові, медичні та охоронні пристрої.

Висновки

Досліджено роботу систем управління жестами як перспективного інтерфейсу людина-машина, що містить наступні етапи: 1) захоплення зображення; 2) локалізація руки; 3) розпізнавання жестів; 4) формування команд. Такі системи дозволяють реалізувати дистанційне управління автоматизованими пристроями не тільки звичайними користувачами, а й людьми з вадами слуху і мови.

Обраний для локалізації руки на зображенні метод сегментації за кольором забезпечує достатню швидкодію для того, щоб його програмна реалізація могла працювати в реальному часі.

Література

1. Абакумов В.Г. Интерпретация движений рук расширяет возможности интерактивного управления в интеллектуальных системах / В.Г. Абакумов, Е.Ю. Ломакина // Природные и интеллектуальные ресурсы Сибири. – 2009. – С.199-202
2. Lomakina O.Y. Gestures Recognition as a New Information Input Device for Automatic System Control / O.Y. Lomakina // «Modern Problems of Radio Engineering, Telecommunications, and Computer Science» : Proceedings of X-th International Conference TCSET'2010. – Lviv-Slavske, 2010. – P. 100-103.
3. Hand tracking and gesture recognition for human-computer interaction / [C. Manresa et al.] // Electronic Letters on Computer Vision and Image Analysis. –2005. – № 5(3). – P. 96-104.
4. Абакумов В.Г. Застосування жестів рук при людино-машинному інтерфейсі / В.Г. Абакумов, О.Ю. Ломакіна, О.Б. Яровенко // Электроника и связь. – Тематический выпуск : Электроника и нанотехнологии. – 2011.
5. Aggarwal J.K. Human Motion Analysis: A Review / Aggarwal J.K. and Q Cai. // Computer Vision and Image Understanding 73. –1999. – № 3. – P. 428-440.
6. Stenger B. (2001) Model-based 3D tracking of an articulated hand / B. Stenger, P.R.S. Mendonca, R. Cipolla // In: The 20th IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'01). – December 2001. – Kauai, HI, US.
7. Kato M. Articulated Hand Tracking by PCA-ICA approach / M. Kato, Y.W. Chen and G. Xu // Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. – Southampton, 2006. – P. 329-333.

Literatura

1. Abakumov E.Ju. Prirodnye I intellektual'nye resursy Sibiri. 2009. S. 199-202.
2. LomakinaO.Y. Proceedings of X-thInternational Conference TCSET'2010, "Modern Problems of Radio Engineering, Telecommunications, and Computer Science". Lviv-Slavske. 2010. P. 100-103.
3. Manresa C. "Hand tracking and gesture recognition for human-computer interaction". Electronic Letters on Computer Vision and Image Analysis.№ 5(3). 2005. P. 96-104.
4. AbakumovV.G. Jelektronika I svjaz'. Tematicheskij vypusk "Jelektronika I nanotehnologii". 2011.
5. Aggarwal J.K. Computer Vision and Image Understanding 73. № 3. 1999. P. 428-440.
6. Stenger B. The 20th IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'01). Kauai. HI. US.December 2001.
7. Kato M. Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. Southampton. 2006. P. 329-330.

Е.Ю. Ломакина
Анализ работы систем распознавания жестов
В приведенной статье рассматривается одно из перспективных направлений развития информационных технологий в области интерфейса человек-машина на основе использования естественного для человека способа общения с компьютерами. Предоставлен обзор основных этапов, процедур и технологий сбора данных, которые применяются в работе системы распознавания жестов. Проводится поэтапный анализ работы систем распознавания жестов, определены основные проблемы их построения и предлагаются оптимальные пути их дальнейшего решения.

E.Y. Lomakina
Analysis of Gesture Recognition System Operation
This article describes one of the most perspective directions of information technology development for human-machine interface based on human natural means of interaction between human and computer. A review of all main stages, procedures and data capture technologies which are used during operation of gesture recognition is represented. The detail analysis of gesture recognition system operation is provided. All main problems of gesture recognition system development and their optimal solutions are defined.
.

Стаття надійшла до редакції 06.06.2011.